Acta virologica **59**: 276 – 283, 2015 doi:10.4149/av_2015_03_276

Analysis of different HCV NS4B domains for the development of global consensus sequence

H. KHAN, S. MURAD*

Molecular Immunology Lab, Health Care Biotech Department, Atta-ur-Rahman School of Applied Biosciences, National University of Sciences and Technology, 44000 Islamabad, Pakistan

Received January 7, 2015; revised February 16, 2015; accepted August 12, 2015

Summary. – The non-structural 4B (NS4B) protein of hepatitis C virus (HCV) is a hydrophobic protein implicated recently in the formation of membranous web, a platform for the formation of replication complex and thus is potential target for antivirals. The CLC main workbench was used to generate genotype-specific consensus sequence, global consensus sequence and a representative phylogenetic tree from non-structural 4 B (NS4B) protein sequences of seven different HCV genotypes reported from all over the world. The C-terminal domain (CTD) of NS4B protein especially the residues involved in interaction with ER membrane were found to be highly conserved. Other residues found to be highly conserved across all HCV genotypes included; 5 aromatic residues of N-terminal domain (NTD) (F49, W50, W55, F57, and Y63), 3 hydrophobic leucine residues (L237, L240, L245), and 2 positively charged residues of CTD (R248 and H250), dimerization motif of transmembrane domain 3 (TMD3) (G_{143}YGAG_{147}) and its surrounding residues (F118 and F155) and TMD1 Ser/Thr cluster residues (T87, S88 and T95) involved in the hydrogen (H) bond interactions. In short, amino acids of NTD, TMD and CTD domains involved in the membrane association/anchoring of NS4B and formation of membranous web are highly conserved and can serve as potential targets for antivirals and peptide vaccines. These conserved residues formed the basis for the development of five short peptides proposed to serve as potential therapeutic target. The phylogenetic analysis was particularly interesting for NS4B sequences of 3a Pakistani isolates. The high degree of variability prevented the clustering of Pakistani isolates with other sequences in phylogenetic tree, revealing geographical disparity.

Keywords: HCV NS4B; global consensus sequence; membranous web; amphipathic alpha-helices; dimerization motif

Introduction

Hepatitis C is an enveloped, positive stranded RNA virus which accounts for 170 million chronic infections leading to hepatitis, fibrosis and even liver cancer (Han *et al*., 2011). Hepatitis C virus (HCV) belongs to the genus *Hepacivirus* of the family *Flaviviridae* (Miller *et al.,* 1990; Simmonds *et al*., 2005). So far, there is no vaccination available for HCV. Due to high mutation rate, there are seven HCV genotypes showing ~30% variability in their genetic sequences (Simmonds *et al*., 2005).The current standard therapies have genotype specific response rates. Most of HCV inhibitors are not designed to have pan-genotypic potential. The development of global consensus sequence against potential therapeutic targets involved in various stages of HCV replication may help to design the specific inhibitors with pan-genotypic effects. Such global consensus sequences have been designed against various HCV proteins such as the glycoproteins, structural and non-structural proteins. No such consensus

*Corresponding author. E-mail: s.mall@asab.nust.edu.pk, sheebamall@yahoo.com; phone: +92-51-9085-6139.

Abbreviations: CTD = C-terminal domain; HCV = hepatitis C virus; MB = membranous web; NS4B = non-structural 4 B; NTD = N-terminal domain; TMD = transmembrane domain

has been established for non-structural protein-4B (NS4B). The HCV genotype-specific effects of NS4B inhibitors such as clemizole and anguizole has been reported. Therefore, the development of a global consensus sequence for NS4B would help to design drugs with a global impact. Therefore, the goal of the current study is to explore the conserved regions of HCV NS4B as a possible drug target.

The HCV viral genome is 9.6 kb long, encoding ten proteins in a single polypeptide chain which is later cleaved to give rise to ten proteins including; three structural proteins (core, E1, and E2), the highly hydrophobic p7 peptide, and six non-structural (NS) proteins (NS2, NS3, NS4A, NS4B, NS5A, and NS5B). The NS proteins are not only involved in HCV replication but some proteins such as; NS2, NS3, and NS5A also contribute to the viral assembly (Yi *et al*., 2007; Ma *et al*., 2008; Masaki *et al*., 2008; Tellinghuisen *et al*., 2008; Han *et al*., 2011). The NS3 acts as both serine protease and helicase (Kim *et al*., 1995; Tackett *et al*., 2001). The NS4A is a cofactor of NS3 and also helps in the binding of NS3 to the host membrane (Wolk *et al*., 2000). The NS5A binds to the HCV RNA and inhibits the action of interferon (Guo *et al*., 2001; Huang *et al*., 2005; Appel *et al*., 2008). The NS5B of HCV is RNA-dependent RNA polymerase which replicates HCV genome in association with other viral and host factors (Bartenschlager and Sparacio 2007; Moradpour *et al*., 2007; Stone *et al*., 2007).

The NS4B of HCV is a 27 K hydrophobic protein found to be an inducer of membranous web (Hugle *et al*., 2001; Moradpour *et al*., 2007). It comprises of an N-terminal domain (NTD) (aa~1 to 79), a transmembrane domain (TMD) harboring four putative transmembrane regions (aa~80 to 191), and a C-terminal domain (CTD) made from up of 70 amino acids (aa~192 to 261) (Lundin *et al*., 2003; Elazar *et al*., 2004; Aligo *et al*., 2009; Gouttenoire *et al*., 2009a,b). The N-terminal part contains two alpha helices involved in NS4B host membrane association (Elazar *et al*., 2004; Gouttenoire *et al*., 2009a). The C-terminal domain (CTD) of NS4B consist of two putative α-helices, represented by residues ~201 to 212 and ~228 to 253, respectively (Gouttenoire *et al*., 2009 b). The CTD is believed to be on the cytosolic side of the endoplasmic reticulum membrane, whereas the location of the NTD remains controversial (Han *et al*., 2011). NS4B is a multi-functional protein that possesses GTPase and ATPase activities (Einav *et al*., 2004) and binds to the 3' end of the negative-sense viral RNA (Einav *et al*., 2008).

Like most of the positive sense RNA viruses, HCV also replicate their genome in association with the intracellular host membranes. These membranes are rearranged by the viruses in the form of novel structures termed as membranous web (Egger *et al*., 2002). Membranous webs have been observed in cells expressing mature NS4B suggesting that NS4B alone induces formation of the membranous webs (Egger *et al*., 2002; Konan *et al*., 2003). Membranous webs are named due to their appearance like membrane vesicles and are believed to be partly derived from the host endoplasmic reticulum. It provides a platform for the viral replication and harbors replication complex including the five replicase proteins (NS3, NS4A, NS4B, NS5A and NS5B) some viral RNAs and host factors (Aligo *et al*., 2009; Manna *et al*., 2010).

Owing to its multi-functionality and role in the induction of membranous webs, NS4B is a new target for the development of antivirals. There is a need to generate a global consensus sequence of NS4B across all HCV genotypes that can serve as potential target for the development of vaccines and antiviral therapies. This can be achieved by aligning available protein sequences of NS4B of different HCV genotypes. Aim of the current study is to develop a global consensus sequence of NS4B protein based on the individual consensus sequences of seven different HCV genotypes and the analysis of different domains of NS4B protein for variability or conservation. Finally, developing a phylogenetic tree would help in analyzing the evolutionary association of NS4B sequences of Pakistani isolates.

Materials and Methods

Retrieval of HCV NS4B sequences. The search for the protein sequences of HCV NS4B from the European HCV database and NCBI databases retrieved 701 sequences (Supplementary Table 1) belonging to seven genotypes and subtypes. These sequences were reported from all over the world including USA, UK, France, China, India, Pakistan, Turkey, South Korea, Australia, Taiwan, Japan, Brazil, Canada, Egypt, Spain, Denmark, Canada etc. 538 sequences were selected for different subtypes of genotype 1 that were reported from USA, China, Ireland, Australia, Denmark, Russia, Turkey, Japan, India, Indonesia and Equatorial Guinea. For different subtypes (2a, 2b, 2c, 2i and 2k) of genotypes 2, 32 NS4B sequences that were reported from Japan, Vietnam, USA and Philippines were selected. Forty three sequences of NS4B for genotype 3 (subtypes 3a, 3g, 3i and 3k) of HCV submitted from Pakistan, India, Japan, United Kingdom and Canada were included in the study. Thirty two sequences of NS4B of HCV genotype 4 from subtypes, 4a, 4b, 4d, 4f, 4g, 4i, 4k, 4m, 4n, 4o, 4p, 4q, 4r, and 4v were selected. These sequences were reported from United Kingdom, Canada, Egypt, USA and France. HCV genotype 5a sequences of NS4B reported from South Africa and China were retrieved. Forty five sequences of genotype 6 subtypes 6a, 6b, 6c, 6d, 6e, 6f, 6g, 6h, 6i, 6j, 6k, 6l, 6m, 6n, 6o, 6p, 6q, 6t with their origin from Thailand, Canada, Vietnam, China, Hong Kong and USA were used in the current study. A single available NS4B sequence of 7a was used as a representative of genotype 7.

Development of consensus sequence. The NS4B protein sequences of all six genotypes were fed to CLC main workbench software to

generate consensus sequence of each genotype. The consensus sequences of six genotypes thus constructed in CLC main workbench software along with single 7a sequence were subjected to alignment in CLC workbench to generate NS4B global consensus sequence of all seven genotypes.

Peptide designing and phylogenetic analysis. Different motifs and domains of NS4B were analyzed to find variation using HCV NS4B global consensus sequence. From HCV NS4B global consensus sequence alignment, short stretches of amino acids were selected from highly conserved region of HCV NS4B. These short peptides could be good targets for potential vaccine testing and development. A representative phylogenetic tree of NS4B protein sequences was constructed using UPGMA method.

Results

Protein consensus sequence of HCV NS4B was generated for each of the seven genotypes using CLC main workbench (Fig. 1). Global consensus sequence of 701 protein sequences of NS4B was developed by aligning the consensus sequences of all six genotypes and a 7a sequence (Fig. 1). The bars represent percent conservation of amino acid residues. The "X" shows highly variable residues whereas conserved amino acid residues are labeled by their symbols. Different regions of global consensus sequences were analyzed for amino acid variability and conservation. Stretches of highly conserved amino acid residues can serve as peptide vaccine. Therefore, short peptides as shown in Table 1 and highlighted in consensus sequence (Fig. 1) were selected from the conserved regions of NS4B. Phylogenetic analysis was done by aligning first all 701 NS4B sequences in CLC workbench followed by drawing phylogenetic tree using UPGMA method. From the phylogenetic tree of 701 NS4B sequences, 210 NS4B representative sequences of different HCV genotypes were selected for the construction of ultimate phylogenetic tree (Supplementary Fig. 1). The excluded NS4B sequences included sequences of same HCV subtype which clustered together and did not show any evolutionary association with NS4B sequences of other HCV genotypes and subtypes. Five NS4B sequences of 3a Pakistani isolates were also aligned in order to analyze the variability and conservation of various residues among them (Fig. 2).

Table 1. Sequences and positions of peptides designed from the highly conserved regions of NS4B

Sequence and position of peptide
S_{59}GIQYLAGLSTLPGNP$_{74}$
F_{118}VVSGLAGAA$_{127}$
G_{150}ALVAFKIM$_{158}$
N_{170}LLPAILSPGALVVGV$_{185}$
Q_{203}WMNRLIAFASRGNHVSPTHY$_{223}$

Discussion

Little is known about the functions of ER-anchored NS4B and phosphorylated form of NS5A, however both are critical players in the genome replication of HCV (Blight, 2011). The NS4B is anchored in the ER membrane through its four transmembrane domains whereas, the N and C-terminal domains are presumably localized towards the cytosolic side of ER membrane (Blight, 2011). The NS4B protein is comprised of three distinct regions; an N-terminal domain (NTD), transmembrane domain (TMD) and a C-terminal domain (CTD) (Lundin *et al*., 2003; Elazar *et al*., 2004; Aligo *et al*., 2009; Gouttenoire *et al*., 2009a,b).

The NTD is 1-79 aa long and contains two amphipathic alpha-helices (AH1 and AH2) (Elazar *et al*., 2004; Gouttenoire *et al*., 2009a). The AH1 (encompassing 6–29 aa) is presumed to interact with ER membrane through its hydrophobic side. Disrupting the helical structure of AH1 is associated with the loss of NS4B ability to form membranous web and replication complex (Elazar *et al*., 2004). The AH2 (42–66 aa long) also mediates ER membrane interaction of NS4B and is critical for the oligomerization of NS4B (Lundin *et al*., 2003; Elazar *et al*., 2004). The global consensus sequence analysis shows that N-terminal domain is comparatively more variable. However, the second AH2 is relatively more conserved highlighting its role in the ER membrane association and oligomerization of NS4B. The last portion of NTD (59–74 aa) has been found to be well conserved among different HCV genotypes. Most of the mutated residues in AH2 and in the last part of NTD exhibit similar physiochemical properties with the wild type which has been reported earlier (Gouttenoire *et al*., 2009a). Moreover, the replacement of the aromatic residues W50, W55, F57, and Y63 of AH2, is believed to play an essential role in ER membrane interface interaction, and together with alanine (A) abrogates formation of membranous web and viral replication (Gouttenoire *et al*., 2009a). These residues were found to have high degree of conservation across different HCV genotypes.

The central region of NS4B consists of four transmembrane domains (TMD1 to TMD4). Besides its functions as ER membrane anchoring of NS4B, the transmembrane domains are implicated in recruitment of viral and host proteins to the replication complex and the formation of NS4B foci/membranous web. The TMD2 and TMD3 contain dimerization motif GlyXXXGly (GXXXG), a common feature of TMD helices (Han *et al*., 2011). Mutation of glycine in dimerization motif such as G125A/L and G129L is associated with *in vitro* suppression of HCV replication (Han *et al*., 2011). Our global consensus sequence shows that all amino acid residues of TMD2 dimerization motif (G_{125}AAVG_{129}) are highly conserved except V128I substitution in genotype 1 and 3. Among the surrounding residues (F118 and V128) of TMD2 dimerization contributing to

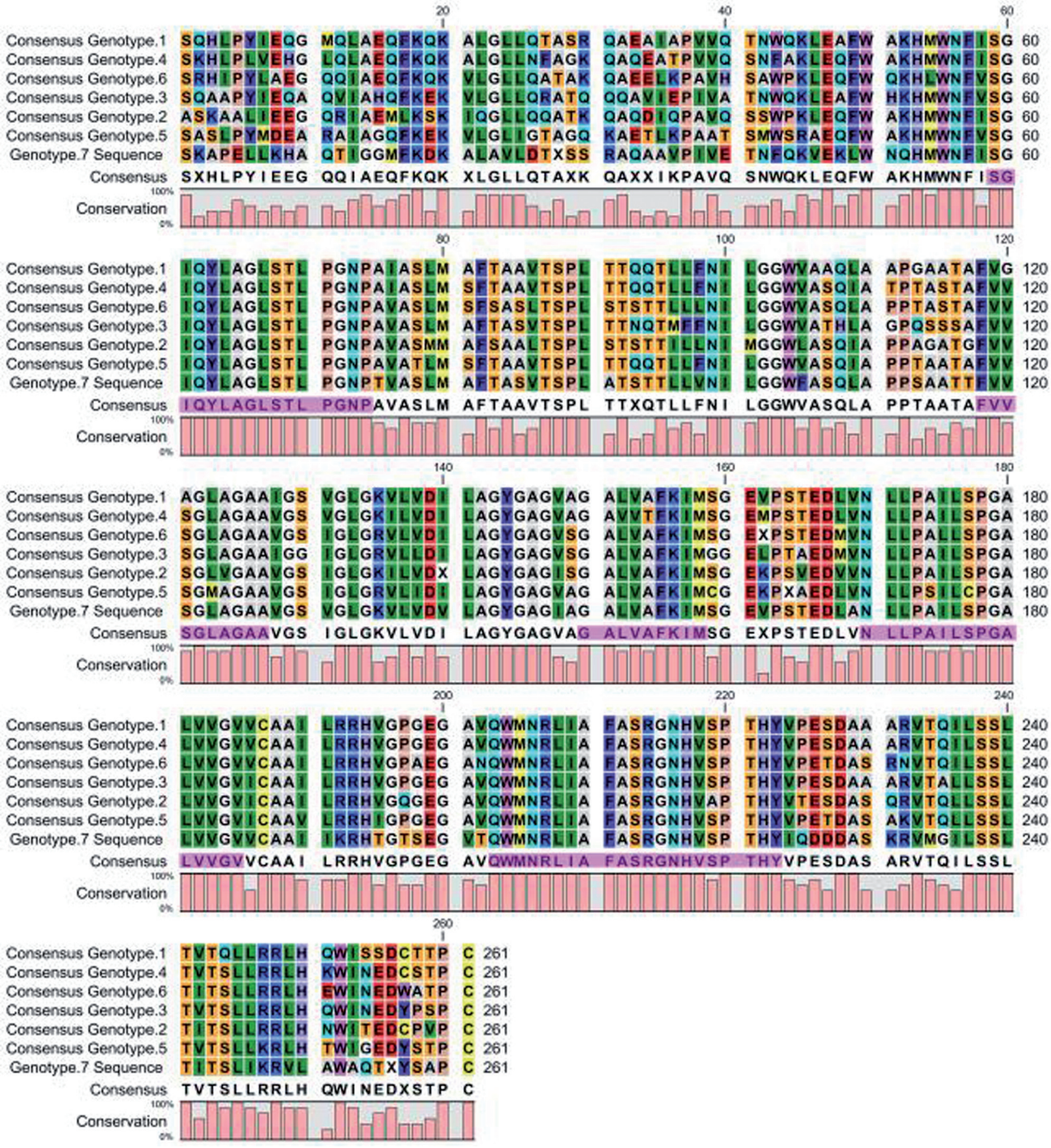


Fig. 1

Multiple sequence alignment of NS4B protein consensus sequences of genotypes 1–6 and a single available NS4B protein sequence of genotype 7a
The global consensus sequence is shown at the base. The conserved residues are represented by their symbols while "X" denotes highly variable residues. Bars under the global consensus sequence show percent degree of amino acid conservation. Stretches of amino acids highlighted in light purple in the global consensus sequence show short peptides.

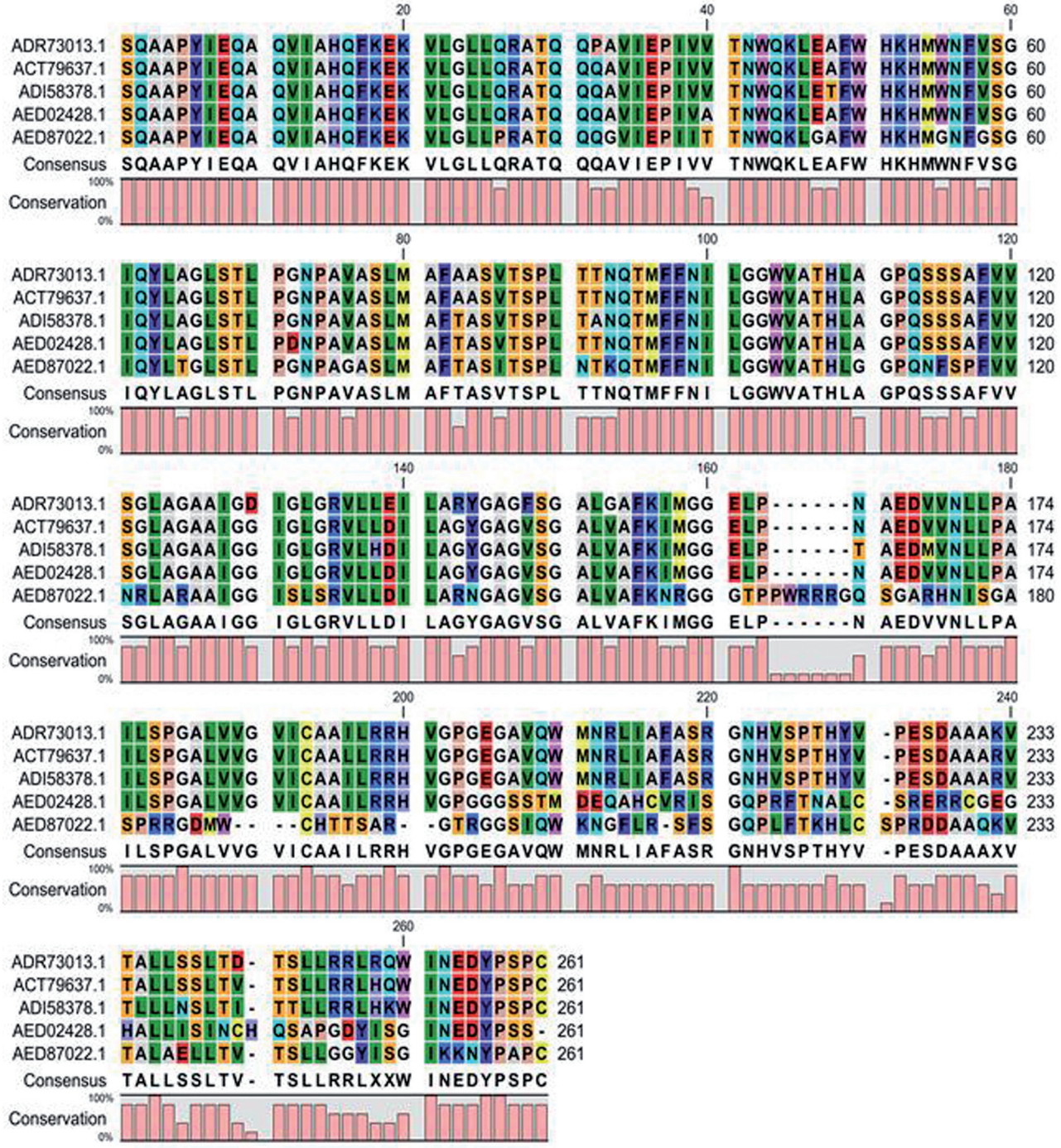


Fig. 2

Multiple sequence alignment of 3a NS4B protein sequences of Pakistani isolates

The consensus sequence is shown at the base. The conserved residues are represented by their symbols while "X" denotes highly variable residues. Bars under the consensus sequence show percent degree of amino acid conservation.

hydrogen bond interaction (Adamian *et al*., 2001; Langosch *et al*., 2009), F118 was found to be highly conserved across all genotypes. However, V128I substitution might support the fact that V128A mutation has negligible effect on the Japanese fulminant hepatitis 1 replication as described previously (Han *et al*., 2011). The dimerization motif $G_{143}YGAG_{147}$ of TMD3 was found to be 100% conserved among all HCV genotypes. The closest F155 residue is also well conserved suggesting its importance in the hydrogen (H) bond interactions in TMD3.

The Ser/Thr (S/T) cluster-like dimerization motif of TMD1 promotes helical interaction through hydrogen bond of hydroxyl group in the side chain of Ser and/or Thr (Dawson *et al*., 2002; Langosch *et al*., 2009). The NS4B consensus sequence alignment reveals that some amino acid residues T87, S88, and T95 of Ser/Thr clusters required for the HCV replication (Han *et al*., 2011) are well conserved. However, T94Q was found be the most prevalent among the serine and threonine residues of Ser/Thr (S/T) cluster-like dimerization motif. The S83A mutation that suppresses HCV replication *in vitro* (Han *et al*., 2011) was not found; however some HCV genotypes harbor S83T substitution.

The C-terminal domain (CTD) of NS4B consist of two putative α-helices, represented by residues ~201 to 212 and ~228 to 253, respectively (Gouttenoire *et al*., 2009 b). The consensus sequence analysis shows that CTD has comparatively high degree of conservation. The hydrophobic leucine (L) residue at positions 237, 240, 246, and 249 has been implicated in an ER membrane interaction of NS4B (Gouttenoire *et al*., 2009 b), protein-protein interaction (Yu *et al*., 2006) or protein-RNA interactions (Egger *et al*., 2002) and is required for viral replication of HCV. Mutation of these leucines (L) to alanine (A) impairs membrane association of NS4B and replication of HCV (Gouttenoire *et al*., 2009b). Our NS4B consensus sequence analysis reveals that L237, L240, and L245 are highly conserved. However, L246I and L249V substitution were found in the NS4B of genotype 7. Both isoleucine and valine like leucine are hydrophobic in nature supporting the fact that these are involved in the vital activities associated with NS4B. The positively charged aa residues K247, R248, and H250 in CTD interact with the negative head groups of lipids of ER membrane and are required for HCV replication (Liefhebber *et al*., 2009). Substitution of these positively charged aa with a negatively charged glutamic acid (K247E/R248E/H250E) severely affect membrane association of NS4B and hence replication of HCV (Liefhebber *et al*., 2009). The consensus sequence analysis shows that these positively charged residues are highly conserved but genotype 7 carries H250L substitution. Yu *et al*. suggested that membrane association of NS4B is mediated by palmitoylation on cysteines 257 and 261 (Yu *et al*., 2006). Our consensus sequence analysis illustrates that residue at 257 is highly variable and represented by "X" in global consensus sequence. However, C261 is well conserved across all HCV genotypes. It has been determined that C257S and C261S substitutions do not influence localization of the NS4B-CTD (Liefhebber *et al*., 2009). This might be an explanation for the variability of cysteine 257.

In the middle of NS4B, there are two nucleotide binding motifs (NBMs), A motif and B motif characterized by conserved elements GXXXXGK and DXXA, respectively (Gorbalenya and Koonin, 1989; Sklan *et al*., 2006). It has been shown that mutations of NBM motifs residues impair GTP binding and hydrolysis ability of NS4B as well as inhibit viral replication (Sklan *et al*., 2006). This suggests that NBMs bind to RNA during HCV replication. The consensus sequence analysis of NS4B elucidates that residue G129 and G134 of motif A ($G_{129}SIGLGK_{135}$) and D228 of motif B ($D_{228}ASA_{231}$) are highly conserved, however, K135 was found to be mutated with a similar positively charged residue, arginine in genotypes 3, 5 and 6. The A231 of motif B was found substituted by three different residues, Q, R, and K in genotypes 3, 6 and 7, respectively.

During the NS4B consensus sequence alignment a rooted phylogenetic tree of 210 NS4B representative sequences was constructed using UPGMA method. The phylogenetic analysis of NS4B shows that root of the tree bifurcated in to two branches, one leads to evolution of 3a (AED87022.1/Pak), second branch bifurcates further and from one wing arises 7a (EF108306) and from second wing evolved genotypes 1, 2, 3, 4, 5, 6 and their subtypes except 3a (AED87022.1/Pak). The phylogenetic analysis of NS4B shows that NS4B sequences of 3a Pakistani isolates do not clustered together and each has an evolutionary association with NS4B sequences of genotype 3 isolates reported from different countries. The phylogenetic analysis suggests that NS4B sequences of 3a isolates from Pakistan have high degree of sequence variability and hence are evolutionary distant from each other.

Our consensus sequence analysis of HCV NS4B suggests that there are certain stretches of amino acids in NTD, TMD and CTD which are involved in membrane association, anchoring, formation of replication complex and viral replication are highly conserved. These stretches of conserved residues can be used for the testing and development of peptide vaccine. Furthermore information about the conserved regions of NS4B could be very helpful for the development of antiviral compounds. The phylogenetic analysis and sequence alignment of NS4B suggests that there is high degree of sequence variability among the NS4B sequences of different 3a isolates reported from Pakistan and evolutionary are not clustered together in the tree.

Supplementary information is available in the online version of the paper.

Acknowledgements. We are obliged to National University of Sciences and Technology (NUST) for providing all kind of help and support.

References

Adamian L, Liang J (2001): Helix-helix packing and interfacial pairwise interactions of residues in membrane proteins. J. Mol. Biol. 311, 891–907. http://dx.doi.org/10.1006/jmbi.2001.4908

Aligo J, Jia S, Manna D, Konan K V (2009): Formation and function of hepatitis C virus replication complexes require residues in the carboxy-terminal domain of NS4B protein. Virol. 393, 68–83. http://dx.doi.org/10.1016/j.virol.2009.07.033

Appel N. Zayas M, Miller S, Krijnse-Locker J, Schaller T, Friebe P, Kallis S, Engel U, Bartenschlager R (2008): Essential role of domain III of nonstructural protein 5A for hepatitis C virus infectious particle assembly. PLoS. Pathog. 4, e1000035. http://dx.doi.org/10.1371/journal.ppat.1000035

Bartenschlager R, Sparacio S (2007): Hepatitis C virus molecular clones and their replication capacity in vivo and in cell culture. Virus. Res. 127, 195–207. http://dx.doi.org/10.1016/j.virusres.2007.02.022

Blight KJ (2011): Charged residues in hepatitis C virus NS4B are critical for multiple NS4B functions in RNA replication. J. Virol. 85, 8158–8171. http://dx.doi.org/10.1128/JVI.00858-11

Dawson JP, Weinger JS, Engelman DM (2002): Motifs of serine and threonine can drive association of transmembrane helices. J. Mol. Biol. 316, 799–805. http://dx.doi.org/10.1006/jmbi.2001.5353

Egger D, Wolk B, Gosert R, Bianchi L, Blum HE, Moradpour D, Bienz K (2002): Expression of hepatitis C virus proteins induces distinct membrane alterations including a candidate viral replication complex. J. Virol. 76, 5974–5984. http://dx.doi.org/10.1128/JVI.76.12.5974-5984.2002

Einav S, Elazar M, Danieli T, Glenn JS (2004): A nucleotide binding motif in hepatitis C virus (HCV) NS4B mediates HCV RNA replication. J. Virol. 78, 11288–11295. http://dx.doi.org/10.1128/JVI.78.20.11288-11295.2004

Einav S, Gerber D, Bryson PD, Sklan EH, Elazar M, Maerkl SJ, Glenn JS, Quake SR (2008): Discovery of a hepatitis C target and its pharmacological inhibitors by microfluidic affinity analysis. Nat. Biotechnol. 26, 1019–1027. http://dx.doi.org/10.1038/nbt.1490

Elazar M, Liu P, Rice CM, Glenn JS (2004): An N-terminal amphipathic helix in hepatitis C virus (HCV) NS4B mediates membrane association, correct localization of replication complex proteins, and HCV RNA replication. J. Virol. 78, 11393–11400. http://dx.doi.org/10.1128/JVI.78.20.11393-11400.2004

Gorbalenya AE, Koonin EV (1989): Viral proteins containing the purine NTP-binding sequence pattern. Nucleic. Acids. Res. 17, 8413–8440. http://dx.doi.org/10.1093/nar/17.21.8413

Gouttenoire J, Castet V, Montserret R, Arora N, Raussens V, Ruysschaert JM, Diesis E, Blum HE, Penin F, Moradpour D (2009a): Identification of a novel determinant for membrane association in hepatitis C virus nonstructural protein 4B. J. Virol. 83, 6257–6268. http://dx.doi.org/10.1128/JVI.02663-08

Gouttenoire J, Montserret R, Kennel A, Penin F, Moradpour D (2009b): An amphipathic alpha-helix at the C terminus of hepatitis C virus nonstructural protein 4B mediates membrane association. J. Virol. 83, 11378–11384. http://dx.doi.org/10.1128/JVI.01122-09

Guo JT, Bichko VV, Seeger C (2001): Effect of alpha interferon on the hepatitis C virus replicon. J. Virol. 75, 8516–8523. http://dx.doi.org/10.1128/JVI.75.18.8516-8523.2001

Han Q, Aligo J, Manna D, Belton K, Chintapalli SV, Hong Y, Patterson RL, van Rossum DB, Konan KV (2011): Conserved GXXXG- and S/T-like motifs in the transmembrane domains of NS4B protein are required for hepatitis C virus replication. J. Virol. 85, 6464–6479. http://dx.doi.org/10.1128/JVI.02298-10

Huang L, Hwang J, Sharma SD, Hargittai MR, Chen Y, Arnold JJ, Raney KD, Cameron CE (2005): Hepatitis C virus nonstructural protein 5A (NS5A) is an RNA-binding protein. J. Biol. Chem. 280, 36417–36428. http://dx.doi.org/10.1074/jbc.M508175200

Hugle T, Fehrmann F, Bieck E, Kohara M, Krausslich HG, Rice CM, Blum HE, Moradpour D (2001): The hepatitis C virus nonstructural protein 4B is an integral endoplasmic reticulum membrane protein. Virol. 284, 70–81. http://dx.doi.org/10.1006/viro.2001.0873

Kim DW, Gwack Y, Han JH, Choe J (1995): C-terminal domain of the hepatitis C virus NS3 protein contains an RNA helicase activity. Biochem. Biophys. Res. Commun. 215, 160–166. http://dx.doi.org/10.1006/bbrc.1995.2447

Konan KV, Giddings TH, Ikeda M, Li K, Lemon SM, Kirkegaard K (2003): Nonstructural protein precursor NS4A/B from hepatitis C virus alters function and ultrastructure of host secretory apparatus. J. Virol. 77, 7843–7855. http://dx.doi.org/10.1128/JVI.77.14.7843-7855.2003

Langosch D, Arkin IT (2009): Interaction and conformational dynamics of membrane-spanning protein helices. Protein. Sci. 18, 1343–1358. http://dx.doi.org/10.1002/pro.154

Liefhebber JM, Brandt BW, Broer R, Spaan WJ, van Leeuwen HC (2009): Hepatitis C virus NS4B carboxy terminal domain is a membrane binding domain. J. Virol. 6, 62. http://dx.doi.org/10.1186/1743-422X-6-62

Lundin M, Monne M, Widell A, Von Heijne G, Persson MA (2003): Topology of the membrane-associated hepatitis C virus protein NS4B. J. Virol. 77, 5428–5438. http://dx.doi.org/10.1128/JVI.77.9.5428-5438.2003

Ma Y, Yates J, Liang Y, Lemon SM, Yi M (2008): NS3 helicase domains involved in infectious intracellular hepatitis C virus particle assembly. J. Virol. 82, 7624–7639. http://dx.doi.org/10.1128/JVI.00724-08

Manna D, Aligo J, Xu C, Park WS, Koc H, Heo WD, Konan KV (2010): Endocytic Rab proteins are required for hepatitis C virus replication complex formation. Virol. 398, 21–37. http://dx.doi.org/10.1016/j.virol.2009.11.034

Masaki T, Suzuki R, Murakami K, Aizak, H, Ishii K, Murayama A, Date T, Matsuura Y, Miyamura T, Wakita T, Suzuki T (2008): Interaction of hepatitis C virus nonstructural protein 5A with core protein is critical for the production of infectious virus particles. J. Virol. 82, 7964–7976. http://dx.doi.org/10.1128/JVI.00826-08

Miller RH, Purcell RH (1990): Hepatitis C virus shares amino acid sequence similarity with pestiviruses and flaviviruses as well as members of two plant virus supergroups. Proc. Natl. Acad. Sci. USA 87, 2057–2061. http://dx.doi.org/10.1073/pnas.87.6.2057

Moradpour D, Penin F, Rice CM (2007): Replication of hepatitis C virus. Nat. Rev. Microbiol. 5, 453–463. http://dx.doi.org/10.1038/nrmicro1645

Simmonds P, Bukh J, Combet C, Deleage G, Enomoto N, Feinstone S, Halfon P, Inchauspe G, Kuiken C, Maertens G, Mizokami M, Murphy DG, Okamoto H, Pawlotsky JM, Penin F, Sablon E, Shin IT, Stuyver LJ, Thiel HJ, Viazov S, Weiner AJ, Widell A (2005): Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes. Hepatol. 42, 962–973. http://dx.doi.org/10.1002/hep.20819

Sklan EH, Glenn JS (2006): HCV NS4B: From Obscurity to Central Stage.

Stone M, Jia S, Heo WD, Meyer T, Konan KV (2007): Participation of rab5, an early endosome protein, in hepatitis C virus RNA replication machinery. J. Virol. 81, 4551–4563. http://dx.doi.org/10.1128/JVI.01366-06

Tackett AJ, Wei L, Cameron CE, Raney KD (2001): Unwinding of nucleic acids by HCV NS3 helicase is sensitive to the structure of the duplex. Nucleic. Acids. Res. 29, 565–572. http://dx.doi.org/10.1093/nar/29.2.565

Tellinghuisen TL, Foss KL, Treadaway J (2008): Regulation of hepatitis C virion production via phosphorylation of the NS5A protein. PLoS. Pathog. 4, e1000032. http://dx.doi.org/10.1371/journal.ppat.1000032

Wolk B, Sansonno D, Krausslic, HG, Dammacco F, Rice CM, Blum HE, Moradpour D (2000): Subcellular localization, stability, and trans-cleavage competence of the hepatitis C virus NS3-NS4A complex expressed in tetracycline-regulated cell lines. J. Virol. 74, 2293–2304. http://dx.doi.org/10.1128/JVI.74.5.2293-2304.2000

Yi M, Ma Y, Yates J, Lemon SM (2007): Compensatory mutations in E1, p7, NS2, and NS3 enhance yields of cell culture-infectious intergenotypic chimeric hepatitis C virus. J. Virol. 81, 629–638. http://dx.doi.org/10.1128/JVI.01890-06

Yu GY, Lee KJ, Gao L, Lai MM (2006): Palmitoylation and polymerization of hepatitis C virus NS4B protein. J. Virol. 80, 6013–6023. http://dx.doi.org/10.1128/JVI.00053-06

Acta virologica **59**: 276 – 283, 2015 doi:10.4149/av_2015_03_276

Supplementary information

Analysis of different HCV NS4B domains for the development of global consensus sequence

H. KHAN, S. MURAD*

Molecular Immunology Lab, Health Care Biotech Department, Atta-ur-Rahman School of Applied Biosciences, National University of Sciences and Technology, 44000 Islamabad, Pakistan

Received January 7, 2015; revised February 16, 2015; accepted August 12, 2015

Supplementary Table 1. List of Acc. Nos. of NS4B sequences included in the multiple sequence analysis

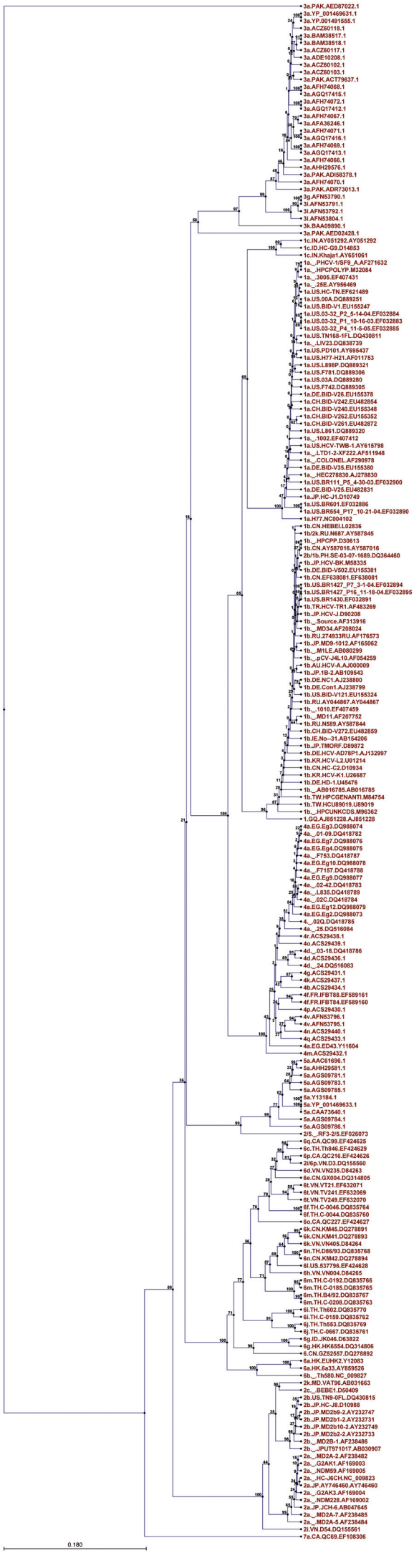
1a. NC004102	1a. AJ851228	1a. EU155345	1a. EU155272	1a. EU155273	1a. EU155274
1a. EU482853	1a. EU155346	1a. EU155347	1a. EU155275	1a. EU482882	1a. EU155311
1a. EU155348	1a. EU482854	1a. EU155349	1a. EU155312	1a. EU155313	1a. EU155314
1a. EU482855	1a. EU482856	1a. EU529681	1a. EU482834	1a. EU155319	1a. EU155320
1a. EU155350	1a. EU155351	1a. EU482872	1a. EU155321	1a. EU482835	1a. EU155322
1a. EU155352	1a. EU482857	1a. EU155353	1a. EU155323	1a. EU155276	1a. EU155277
1a. EU155354	1a. EU155355	1a. EU482858	1a. EU482843	1a. EU482844	1a. EU155278
1a. EU482831	1a. EU155378	1a. EU482873	1a. EU482845	1a. EU482846	1a. EU155233
1a. EU155379	1a. EU482832	1a. EU155380	1a. EU482861	1a. EU155236	1a. EU482862
1a. D10749	1a. DQ889251	1a. DQ889252	1a. EU482863	1a. EU155237	1a. EU482864
1a. DQ889253	1a. DQ889254	1a. DQ889255	1a. EU482865	1a. EU482866	1a. EU482867
1a. DQ889256	1a. DQ889257	1a. DQ889258	1a. EU155238	1a. EU482847	1a. EU482836
1a. DQ889259	1a. DQ889260	1a. DQ889261	1a. EU250017	1a. EU529677	1a. EU155239
1a. DQ889262	1a. DQ889263	1a. DQ889264	1a. EU529678	1a. EU482887	1a. EU155240
1a. DQ889265	1a. DQ889266	1a. DQ889267	1a. EU482868	1a. EU155282	1a. EU155283
1a. DQ889268	1a. DQ889269	1a. DQ889270	1a. EU482869	1a. EU155241	1a. EU155242
1a. DQ889271	1a. DQ889272	1a. DQ889273	1a. EU482870	1a. EU529679	1a. EU482871
1a. DQ889274	1a. DQ889275	1a. DQ889276	1a. EU529680	1a. EU155243	1a. EU482848
1a. DQ889277	1a. DQ889278	1a. DQ889279	1a .EU155244	1a. EU155245	1a. EU155246
1a. EF032883	1a. EF032884	1a. EF032885	1a. EF032900	1a. EF032895	1a. EF032891
1a. DQ889280	1a. DQ889281	1a. DQ889282	1a. EF032890	1a. EF032886	1a. DQ889301
1a. DQ889283	1a. DQ889284	1a. DQ889285	1a. DQ889302	1a. DQ889303	1a. DQ889304
1a. DQ889286	1a. DQ889287	1a. DQ889288	1a. DQ889305	1a. DQ889306	1a. AF011753
1a. DQ889289	1a. DQ889290	1a. DQ889291	1a. EF621489	1a. AY615798	1a. DQ889307
1a. DQ889292	1a. DQ889293	1a. DQ889294	1a. DQ889308	1a. DQ889309	1a. DQ889310
1a. DQ889295	1a. DQ889296	1a. DQ889297	1a. DQ889311	1a. DQ889312	1a. DQ889313
1a. DQ889298	1a. DQ889299	1a. DQ889300	1a. DQ889314	1a. DQ889315	1a. DQ889316

Supplementary Table 1. (continued)

1a. EU155247	1a. EU260395	1a.EU155213	1a. DQ889317	1a. DQ889318	1a. DQ889319
1a. EU155214	1a. EU155215	1a. EU260396	1a. DQ889320	1a. DQ889321	1a. AY695437
1a. EU155216	1a. EU155338	1a. EU482837	1a. DQ430811	1a. EF407412	1a. EF407457
1a. EU155339	1a. EU482889	1a. EU234064	1a. EF407432	1a. EF407428	1a. EF407437
1a. EU155340	1a. EU482850	1a. EU482878	1a. EF407441	1a. EF407452	1a. EF407434
1a. EU155341	1a. EU155342	1a. EU155343	1a. EF407449	1a. AY956468	1a. AY956464
1a. EU155344	1a. EU482852	1a. EU155284	1a. AY956469	1a. EF407431	1a. EF407415
1a. EU155285	1a. EU155286	1a. EU155287	1a. EF407416	1a. EF407456	1a. EF407453
1a. EU155288	1a. EU155289	1a. EU155290	1a. EF407417	1a. EF407419	1a. EF407445
1a. EU155291	1a. EU155292	1a. EU155293	1a. EF407413	1a. EF407435	1a. EF407411
1a. EU155294	1a .EU482876	1a. EU155295	1a. EF407424	1a. AY956466	1a. AY956463
1a. EU155296	1a. EU155297	1a. EU155298	1a. AY956465	1a. EF407422	1a. EF407448
1a. EU155299	1a. EU155248	1a. EU155249	1a. EF407450	1a. EF407427	1a. EF407444
1a. EU155309	1a. EU155310	1a. EU155250	1a. EF407454	1a. EF407425	1a. EF407440
1a. EU155251	1a. EU482838	1a. EU155252	1a. EF407430	1a. EF407438	1a. EF407414
1a. EU482884	1a. EU482840	1a. EU482841	1a. EF407418	1a. EF407443	1a. EF407420
1a. EU155265	1a. EU482842	1a. EU529676	1a. EF407436	1a. EF407451	1a. EF407429
1a. EU155266	1a. EU155267	1a. EU155268	1a. EF407433	1a. EF407455	1a. EF407439
1a. EU155269	1a. EU155270	1a. EU155271	1a. EF407447	1a. EF407446	1a. EF407442
1a. EF407423	1a. EF407421	1a. EF407426	1b. EU155258	1b. EU482879	1b. EU482880
1a. AF290978	1a. AJ278830	1a. M32084	1b. EU155259	1b. EU155260	1b. EU155261
1a. DQ838739	1a. AF511948	1a. AF511949	1b. EU155262	1b. EU482881	1b. EU155263
1a. AF511950	1a. AF271632	1b. AJ000009	1b. EU155264	1b. EU155315	1b. EU155316
1b. EU482859	1b. EU155356	1b. EU155357	1b. EU155317	1b. EU155318	1b. EU155279
1b. EU155358	1b. EU155359	1b. EU155360	1b. EU155280	1b. EU155234	1b. EU482885
1b. EU155361	1b. EU155362	1b. EU155363	1b. EU155235	1b. EU482886	1b. EU155281
1b. EU482874	1b. EU155364	1b. EU155365	1b. EF032894	1b .EF407459	1b. EF407473
1b. EU529682	1b. EU155366	1b. EU155367	1b. EF407481	1b. EF407467	1b. EF407458
1b. EU155368	1b. EU155369	1b. EU155370	1b. EF407495	1b. EF407500	1b. EF407502
1b. EU155371	1b. EU482860	1b. EU155372	1b. EF407491	1b. EF407474	1b. EF407475
1b. EU155373	1b. EU155374	1b. EU155375	1b. EF407476	1b. EF407493	1b. EF407478
1b. EU482875	1b. EU155376	1b. EU155377	1b. EF407480	1b. EF407468	1b. EF407496
1b. AY587016	1b. EF638081	1b. D10934	1b. EF407498	1b. EF407466	1b. EF407469
1b. AY460204	1b. L02836	1b. EU155381	1b. EF407486	1b. EF407472	1b. EF407479
1b. EU482833	1b. EU155382	1b. AJ238799	1b. EF407483	1b. EF407470	1b. EF407497
1b. AJ132997	1b. U45476	1b. AJ238800	1b. EF407461	1b. EF407471	1b. EF407463
1b. AB154206	1b. AB109543	1b. M58335	1b. EF407484	1b. EF407503	1b. EF407465
1b. D90208	1b. D50480	1b. D50481	1b. EF407462	1b. EF407482	1b. EF407487
1b. D50484	1b. AB426117	1b. AF139594	1b. EF407490	1b. EF407499	1b. EF407489
1b. D63857	1b. AB191333	1b. D16435	1b. EF407492	1b. EF407485	1b. EF407501
1b. D14484	1b. X61596	1b. D11355	1b. EF407494	1b. EF407460	1b. EF407488
1b. DQ244140	1b. DQ244141	1b. AF165046	1b. EF407477	1b. EF407464	1b. EF407504
1b. AF165064	1b. AF165048	1b. AF165050	1b. AB016785	1b. D89815	1b. AY045702
1b. AF165052	1b. AF165054	1b. AF165056	1b. AF333324	1b. D85516	1b. AF356827
1b. AF165058	1b. AF165060	1b. AF165062	1b. AB049087	1b. AB049088	1b. AB049089
1b. D89872	1b. U26687	1b. U01214	1b. AB049090	1b. AB049091	1b. AB049092

Supplementary Table 1. (continued)

1b. AF176573	1b. AY044867	1b. AY587844	1b. AB049093	1b. AB049094	1b. AB049095
1b. AF483269	1b. U89019	1b. M84754	1b. AB049096	1b. AB049097	1b. AB049098
1b. EU155324	1b. EU155325	1b. EU482888	1b. AB049099	1b. AB049100	1b. AB049101
1b. EU155326	1b. EU155327	1b. EU234061	1b. D45172	1b. D30613	1b. .M96362
1b. EU155328	1b. EU155329	1b. EU155330	1b. AB080299	1b. AF207752	1b. AF207753
1b. EU155331	1b. EU155332	1b.EU155333	1b. AF207754	1b. AF207755	1b. AF207756
1b. EU155334	1b. EU234062	1b. EU155335	1b. AF207757	1b. AF207758	1b. AF207759
1b. EU482849	1b. EU155336	1b. EU155337	1b. AF207760	1b. AF207761	1b. AF207762
1b. EU155217	1b. EU155218	1b. EU155219	1b. AF207763	1b. AF207764	1b. AF207765
1b. EU155220	1b. EU155221	1b. EU155222	1b. AF207766	1b. AF207767	1b. AF207768
1b. EU155223	1b..EU155224	1b. EU482883	1b. AF207769	1b. .AF207770	1b. AF207771
1b. EU155225	1b. EU155226	1b. EU155227	1b. AF207772	1b. AF207773	1b. AF207774
1b. EU155228	1b. EU155229	1b. EU155230	1b. AF208024	1b. AF054259	1b. AF313916
1b. EU155231	1b. EU155232	1b. EU155300	1b/2k. AY587845	1c. D14853	1c. AY051292
1b. EU482877	1b. EU155301	1b. EU155302	1c. AY651061	2/5. EF026073	2a. AY746460
1b. EU155303	1b. EU155304	1b. EU155305	2a. AB047645	2a. .AF169003	2a. AF169004
1b. EU155306	1b. EU155307	1b. EU155308	2a. 009823	2a. AF238481	2a. AF238482
1b. EU155253	1b. EU155254	1b. EU155255	2a. AF238483	2a. AF238484	2a. AF238485
1b. EU155256	1b. EU155257	1b. EU482839	2a. AF169002	2a. AF169005	2b. D10988
2b. AY232731	2b. AY232749	2b. AY232733	2b. AY232741	2b. AY232743	2b. AY232745
2b. AY232735	2b. AY232737	2b. AY232739	2b. AY232747	2b. DQ430815	2b. AB030907
2b. AF238486	2b/1b. Q364460	2c. D50409	2i. DQ155561	2i/6p. DQ155560	2k. AB031663
3a.YP.001491555.1	3a.AED87022.1	3a.ADR73013.1	3a.ACZ60107.1	3a.ACZ60106.1	3a.ACZ60105.1
3a.ADI58378.1	3a.ACT79637.1	3a.AED02428.1	3a.ACZ60104.1	3a.ACZ60103.1	3a.ACZ60102.1
3a.YP_001469631.1	3a.AGQ17416.1	3a.AGQ17415.1	3a.ADE10208.1	3a.AFH74072.1	3a.AFH74071.1
3a.AGQ17413.1	3a.AGQ17412.1	3a.AHH29576.1	3a.AFH74070.1	3a.AFH74069.1	3a.AFH74068.1
3a.BAM38518.1	3a.BAM38517.1	3a.AFA36246.1	3a.AFH74067.1	3a.AFH74066.1	3i.AFN53804.1
3a.ACZ60118.1	3a.ACZ60117.1	3a.ACZ60114.1	3a.ACZ60109.1	3a.ACZ60108.1	3g.AFN53790.1
3a.ACZ60113.1	3a.ACZ60112.1	3a.ACZ60111.1	3k.BAA09890.1	3i.AFN53792.1	3i.AFN53791.1
3a.ACZ60110.1	4n.ACS29440.1	4o.ACS29439.1	4r.ACS29438.1	4k.ACS29437.1	4d.ACS29436.1
4b.ACS29434.1	4q.ACS29433.1	4m.ACS29432.1	4a. DQ988077	4a. DQ418782	4a. DQ418783
4g.ACS29431.1	4p.ACS29430.1	4v.AFN53795.1	4a. DQ418784	4a. DQ516084	4a. DQ418788
4v.AFN53796.1	4. .DQ418785	4a. Y11604	4a. DQ418787	4a. .DQ418789	4d. DQ418786
4a. DQ988078	4a. DQ988079	4a. DQ988073	4d. DQ516083	4f. EF589160	4f. EF589161
4a. DQ988074	4a. DQ988075	4a. DQ988076	5a.YP_001469633.1	5a.AHH29581.1	5a.Y13184.1
5a.CAA73640.1	5a.AAC61696.1	5a.AGS09785.1	5a.AGS09786.1	5a.AGS09783.1	5a.AGS09784.1
5a.AGS09781.1	6. DQ278892	6a. AY859526	6a. DQ480514	6a. DQ480515	6a. DQ480518
6a. DQ480513	6a. DQ480516	6a. DQ480523	6a. DQ480519	6a. DQ480520	6a. DQ480521
6g. D63822	6h. D84265	6i. DQ835762	6a. DQ480522	6a. DQ480517	6a. DQ480524
6i. DQ835770	6j. DQ835761	6j. DQ835769	6a. DQ480512	6a. Y12083	6b. NC_009827
6k. DQ278893	6k. DQ278891	6k. D84264	6c. EF424629	6d. D84263	6e. DQ314805
6l. EF424628	6m. DQ835767	6m. DQ835765	6f. DQ835760	6f. DQ835764	6g. DQ314806
6m. DQ835766	6m. DQ835763	6n. DQ278894	6n. DQ835768	6o. EF424627	6p. EF424626
6t. EF632071	6q. EF424625	6t. .EF632069	6t. EF632070	7a. EF108306	



Supplementary Fig. 1

Phylogenetic tree of 210 representative HCV NS4B protein sequences of seven HCV genotypes